

REMARKS

Claims 1-66 are all the claims pending in the application. Claims 1-5, 14-27, 36-49, and 58-66 stand rejected on prior art grounds. The Office Action indicates that claims 6-13, 28-35, and 50-57 are objected to only as being dependent on a rejected base claim. Claims 1, 3, 4, 6, 23, 25, 26, 28, 45, 47, 48, and 50 are amended herein. Moreover, the specification is amended to correct a typographical error. No new matter is being added. Applicants respectfully traverse the objections/rejections based on the following discussion.

I. The Prior Art Rejections

Claims 1-5, 14-27, 36-49 and 58-66 stand rejected under 35 U.S.C. §102(e) as being anticipated by Nakao et al. (US Patent No. 6,141,443), hereinafter referred to as Nakao.

Applicants respectfully traverse these rejections based on the following discussion.

Nakao teaches a character extraction apparatus that provides for extracting character data for each character from a text image which is represented by first pixels corresponding to character images and second pixels corresponding to background images. The character extraction apparatus comprises a character row detecting means for detecting character rows from the text image and obtaining position data of each character row; a pixel array extracting means for extracting arrays of continuous first pixels in an area specified by the character row position data and computing position data of each of the arrays of continuous first pixels; a character array linking means for linking the arrays of continuous first pixels in the area based on the position data of the arrays of continuous first pixels; and a character extracting means for

recognizing each set of arrays of continuous first pixels linked by the character array linking means as a character and outputting character data.

However, the amended claims include elements not taught or suggested in Nakao. For example, Nakao does not teach "[a] method for clustering data points with defined quantified relation values between them comprising obtaining a lead value for each data point, wherein said lead value for each data point is derived by taking a sum of all relation values input into said data point plus a frequency associated with said data point, ranking each data point in a lead value sequence list in descending order of lead value, assigning a first data point in said lead value sequence list as a leader of a first cluster, considering each subsequent data point in said lead value sequence list as a leader of a new cluster if its relationship with leaders of each of the previous clusters is less than a defined threshold value or as a member of at least one cluster where its relationship with a cluster leader is at least equal to said threshold value, wherein the threshold value is adaptively found for a given number of clusters, and generating a text summarization of any of a single document and a collection of documents by segmenting a given text input comprising said data points into clusters, and forming a set of leaders of said clusters to represent said text summarization" as claimed in amended independent claim 1, and similarly recited in independent claims 23 and 45.

The Office Action contends that the lead values of the claimed invention are analogous to the similarity values (see Figure 43 of Nakao). However, this is an erroneous assumption because in Nakao, each character is assigned a lead value, which is the same as its similarity. This lead value depends on the character; cluster pair. Whereas, the lead value in the claimed invention is defined for every data point and are not the same as the similarity. In fact, the lead

values in the claimed invention may depend on the data point's similarity with other data points, and not just its own as is the case in Nakao. Furthermore, even if it were assumed that the similarity values in Nakao are analogous to the lead values in the claimed invention, the manner of calculating the lead values in each case is quite different. Col. 32, lines 10-19 of Nakao, which refers to Formula 1 in col. 18 of Nakao generally describe that the similarity may be determined by taking the average (center) between the data points and then divides this average value by some arbitrary constant, C. Conversely, the lead values in the claimed invention are derived by taking a sum of all relation values input into the data point plus a frequency associated with the data point. This is a completely different approach than the approach described in Nakao.

This is best demonstrated in Example 1 given on pages 16-18 of the Applicants' specification as well as the Applicants' Figure 1. In Figure 1, for data point d_1 the lead value is determined by summing all of the inputs into d_1 plus a frequency associated with d_1 (wherein the frequency equals 1.0). Thus, the lead value for $d_1 = 0.25 + 1.0 = 1.25$, similarly for data point d_2 the lead value is determined by summing all of the inputs into d_2 ($0.85 + 0.25 + 0.8$) + $1.0(\text{frequency}) = 2.9$. The remaining data points in Figure 1 are similarly calculated. As such, the method of determining the lead value in the claimed invention is wholly unique and patentably distinct from the calculations in Nakao, assuming the similarity functions in Nakao are analogous to the lead value calculations in the claimed invention, which again, the Applicants contend are not analogous.

Therefore, the amended claims differ from Nakao, which relates only to character recognition. As previously mentioned, the claimed invention relates to unsupervised learning;

namely clustering a given set of data points. Conversely, Nakao deals with supervised learning; namely given a set of labeled data points, the system would predict the label of a data point whose label is unknown.

In Nakao, each character is compared with labeled ("R", "P", "F", etc in the example) clusters in similarity. Based on this similarity each character is assigned to an appropriate cluster. Before Nakao computes the similarity, the labeled clusters are formed from the labeled data points (images representing a particular letter in the example). Whereas, in the claimed invention, the Applicants identify a new way of clustering unlabeled data points, wherein the purpose of finding the similarity is different in the claimed invention compared with Nakao.

In view of the foregoing, the Applicants respectfully submit that the features defined by amended independent claims 1, 23, and 45 contain patentable subject matter and as such, claims 1, 23, and 45 are patentable. Further, dependent claims 2-22, 24-44, and 46-66 are similarly patentable not only by virtue of their dependency from patentable independent claims, respectively, but also by virtue of the additional features of the invention they define. Thus, the Applicants respectfully request that these rejections be reconsidered and withdrawn. Moreover, the Applicants note that all claims are properly supported in the specification and accompanying drawings, and no new matter is being added. In view of the foregoing, the Examiner is respectfully requested to reconsider and withdraw the rejections.

II. Formal Matters and Conclusion

With respect to the objections/rejections to the claims, the claims have been amended, above, to overcome these objections/rejections. Moreover, the specification has been amended

to correct an obvious typographical error. In view of the foregoing, Applicants submit that claims 1-66, all the claims presently pending in the application, are patentably distinct from the prior art of record and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest possible time.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary. Please charge any deficiencies and credit any overpayments to Attorney's Deposit Account Number 09-0441.

Respectfully submitted,

Dated: May 28, 2004



Mohammad S. Rahman
Reg. No. 43,029
McGinn & Gibb, P.L.L.C.
2568-A Riva Road, Suite 304
Annapolis, MD 21401
Voice: (301) 261-8625
Fax: (301) 261-8825
Customer Number: 29154